DOCUMENT RESUME

ED 246 078 TM 840 349

AUTHOR Stiles, Richard L.; And Others
TITLE CASAS: An Effective Measurement System for Life Skills.
PUB DATE Apr 84
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984) and the National Council on Measurement in Education (New Orleans, LA, April 26, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adult Basic Education; Adult Education; Adult Students; *Competency Based Education; Daily Living Skills; High Schools; Item Banks; State Programs
IDENTIFIERS *California Adult Student Assessment System

ABSTRACT

The California Adult Student Assessment System (CASAS) is a comprehensive educational system designed to enable adult educators to develop and evaluate a life skills curriculum for competency based educational programs. The system comprises the CASAS Competency List, the CASAS Item Bank, the User's Manual, the Curriculum Index and Matrix, and Implementation Workshops. The CASAS Item Bank is referenced to the Competency List and consists of more than 2,400 test items which measure life skills. The User's Manual provides technical information related to test construction and field test studies. CASAS has the potential for increasing the precision in placement, monitoring, certification, and evaluation of students and programs at local, state, and national levels. Validity and reliability of items are discussed. The long term effects of CASAS implementation are increased accountability of literacy programs and meaningful reporting of progress. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED246078

CASAS:* An Effective Measurement System for Life Skills

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. L. Stiles

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Organizer: John N. Davis, San Jose Unified Schools
Moderator: Paul Williams, Maryland Department of Education

Participants:
Richard L. Stiles, California Department of Education
John Wise, Elsinore (CA) Union High School District
Patricia L. Rickard, San Diego Community College District
John N. Davis, San Jose Unified Schools

Discussant: Paul Williams, Maryland Department of Education

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 830 679

A Symposium Presented at National Council on Measurement in Education, New Orleans, LA, April 26, 1984

* California Adult Student Assessment System

CASAS: An Effective Measurement System for Life Skills

The California Adult Student Assessment System (CASAS) is a comprehensive educational system designed by field practitioners in order to provide adult education agencies with effective assessment materials and procedures. This system is adaptable to a variety of educational settings, and is linked to competency-based curriculum goals and instructional programs. CASAS can assist adult education agencies with many different needs, ranging from the implementation of life skills literacy curricula, placement of adult students, and assessment of student growth.

History and Need

In 1978, Richard Stiles, measurement and evaluation specialist for the California State Department of Education, conducted a formal, statewide assessment of adult basic education (ABE) programs in California. During that review, Stiles discovered that California adult basic education programs had no uniform standards. They were characterized as having high student dropout rates, traditional school-age curricular material, and ineffective student assessment procedures. Based on these findings, California State Department of Education decided to institute new standards and procedures for the design and certification of adult basic education programs. Exploratory research and investigation of the field of adult education led the state to select a competency-based, life skills literacy approach to ABE instruction. In 1981, the three-year State Plan for Adult Basic Education was adopted to reflect California's commitment to a life skills literacy program. In that Plan, the state mandated the establishment of instructional programs founded upon the principles of competency-based education (CBE). In the Plan, funded agencies were allowed a three-year phase-in of the competency-based management system.

In response to these mandates from the California State Department of Education, three major state-wide projects were instituted to help districts implement competency-based approaches to adult basic education. One of these three projects was the CASAS Consortium, a group of adult education practitioners from districts and agencies both within and outside of California. Specifically, the consortium includes representatives from 40 districts and agencies in California, and also representatives from Washington, Oregon, Idaho, Hawaii, Nevada, Arizona and New York City. The consortium worked together to develop a comprehensive assessment system for improving competency-based instruction and management in adult basic education programs.

In terms of student characteristics, the CASAS programs include the full range of adults and secondary students who are participating in ABE, ESL and high school completion programs in California. That is, all races and ethnicities are enrolled in these programs, including Black, Hispanic, Native Americans, Vietnamese, Chinese, Laotian, Korean, and Caucasian students. Furthermore, the students are of various ages (ranging from 16- to 18-year olds to senior citizens) and low income levels. Program participants include native English speakers, as well as individuals whose primary language is other than English. Participants are typically the least educated, most-in-need adults in California.

Theoretical/Conceptual Bases of Program Approach

The State of California has made a long-term commitment to the development of life skills literacy programs. CASAS is one of three major state-wide projects designed to support the mandated establishment of instructional programs founded upon the principles of competency-based education (CBE). The CBE model is viewed as the most appropriate vehicle to accomplish the primary purpose of the Adult Education Act: "To enable all adults to acquire basic skills necessary to function in society." The success of the project has been evaluated as the degree to which CASAS implementation has moved agencies toward CBE instruction, which is characterized by assessment, instruction and certification for skills needed to function in society.

The development of CASAS fills a critical void in the measurement of literacy for adult and secondary students. The most frequently used literacy tests are inappropriate for adult/secondary students.

In a final report to the State Department of Education in 1977 from one large ABE program in this state, this recommendation was included: "The necessity of evaluating adults by children's grade levels as must be done with the TABE (Tests of Adult Basic Education) and ABLE (Adult Basic Learning Examination) must be eliminated." It is almost impossible to equate a person's communication skills with grade level. The final report in 1978 reflected the same message. "The grade level gain does not accurately reflect student growth. Program emphasis, especially with ESL students, was on the acquisition of competencies needed for survival in everyday life situations. The excellent growth achieved in this area cannot be measured effectively in grade level placement." A similar report issued in the Summer of 1977 stated, "Eliminate the necessity for reporting average growth in grade placement levels. It is not compatible with competency-based education."

This description of why CASAS was developed helps to understand a caveat relating to its use; the assessment system cannot be used successfully in traditional, time-based, course/grade-oriented, credit or seat-time driven programs. Because these traditional types of programs typically measure achievement against grade-level norms established with children, comparisons with CBE programs along dimensions of student achievement are meaningless.

CASAS was designed to impact at the institutional level, making possible articulation among program levels and establishing a uniform method for reporting progress. The long-term effect of widespread CASAS implementation is increased accountability of literacy programs and meaningful reporting of progress at local, state and national levels.

Institutional Self-Assessment Measure

The Institutional Self-Assessment Measure (ISAM) was used to assess the degree of CBE implementation for each agency. The instrument includes 89 items with corresponding 5-point behaviorally anchored rating scales which are used to measure the extent to which CBE minimum program elements have been implemented. The measure yields a total CBE implementation profile score, as well as sub-scores for management, instruction and guidance strands. To measure the effects of CASAS, the total profile score and the sub-scores for management and instruction were used, as these test strands most closely relate to the intervention.

The ISAM is based upon minimum program elements for CBE implementation identified in the California ABE three-year State Plan. Items and scoring criteria were reviewed for validity by a state-wide staff development consortium and refined for field testing with 25 agencies by June, 1983. Interrater reliability was computed at 95% agreement for scale ratings of four agencies by three independent raters. The measure represents a slight modification of formats generally used for Behavioral Observation Scales (BOS) and Behaviorially Anchored Rating Scales (BARS). The measure was carefully constructed to conform with specifications of rating quality reported by Bernardin (1977), Dunnette & Borman (1979), Landy & Farr (1980), and Jacobs, Kafry and Zedeck (1980). The State of California has adopted the ISAM to measure compliance and certify agencies that offer CBE instruction.

Overall Goals/Objectives

A comprehensive educational assessment system for adults has been developed with the goal of assisting adult education agencies in implementing an assessment system designed to reflect competency-based curricula occurring in local instructional programs. The project has achieved four major goals:

- To provide effective assessment materials and procedures for correct placement of adult students into adult programs from beginning through advanced levels of ABE and ESL as well as high school completion programs
- To provide effective assessment for measuring student achievement in a competency-based adult education program
- To provide effective assessment for certifying attainment of competencies required for local certification programs
- To provide effective assessment materials and procedures that are adaptable to a variety of educational settings and are linked to the competency-based curriculum goals and instructional programs

Salient Features of CASAS

The California Adult Student Assessment System is a comprehensive set of procedures and resources designed to enable adult educators to develop and evaluate a life skills curriculum for competency-based educational programs. CASAS provides an inter-related, flexible system that links curriculum, instruction and assessment, allows for measurement of growth through program levels, and provides a means of assessing students with tests designed to measure those competency statements selected and taught at the various difficulty levels. The system is comprised of four major elements which are described below:

CASAS Competency List. This list contains over 100 competency statements within the general life skills content areas of Health, Occupational Knowledge, Government and Law, Consumer Economics and Community Resources. Sources for the content areas and competency statements included in this list were adult education projects validated by the JDRP; namely, the Adult Performance Level (APL) project, the External Diploma Program, and the CLASS Project.

CASAS Item Bank and User's Manual. The CASAS Consortium has created an item bank of more than 2,400 test items that have been developed, field tested, and calibrated to measure life skills along a continuum of difficulty. The test items are referenced to the Competency List, and each item is coded to specify the content area, competency statement and task being measured. In addition to the test items, which can be used to develop customized tests for various measurement purposes, there is a scaling procedure built into the item bank that shows the probability of a correct response at any point along the continuum of content used. The bank of calibrated items is based on the work of Rasch (Rasch 1960, Wright 1968, and Wright and Stone 1979), which is being utilized for item and test development because of its superior psychometric properties (Haladyna and Roid 1983, Hambleton 1983, Kolen and Whitney 1982, and Rentz 1981). Tests that are developed employing this latent trait scaling procedure can provide more relevant information for test development than the use of classical techniques alone.

The CASAS Scale was anchored to intermediate level students' performance on a field test form. This field test form served as a reference point or base from which to determine the relative difficulty of all other test forms. A comparison of forms is possible through an item block design linking items from this base test to all other field test forms. Thus, the entire Item Bank is calibrated on a common interval scale. From this scale, three levels of survey achievement tests were constructed to measure student progress across the state in beginning through advanced levels of ABE and ESL. Two equivalent test forms were constructed for each program level. Content and range of item difficulty were matched in designing the equivalent forms. An achievement scale was calculated for each test form. This scale permits measurement of student progress on an equal interval scale, independent of norm groups. The test also provides criterion/content information about student mastery, since every item has a scaled standardized difficulty level and also is referenced to each specific competency statement selected. Test results can then be used to provide more valid data on individual student achievement as well as group achievement after an instructional intervention, such as 100 hours of instruction.

The User's Manual, which accompanies the Item Bank, provides technical information related to test construction and the field test studies conducted to establish the validity and reliability of the assessment system. In terms of test construction, CASAS has utilized the professional expertise of external measurement and evaluation specialists over the entire four years of development. Content validity was established through a panel of adult education

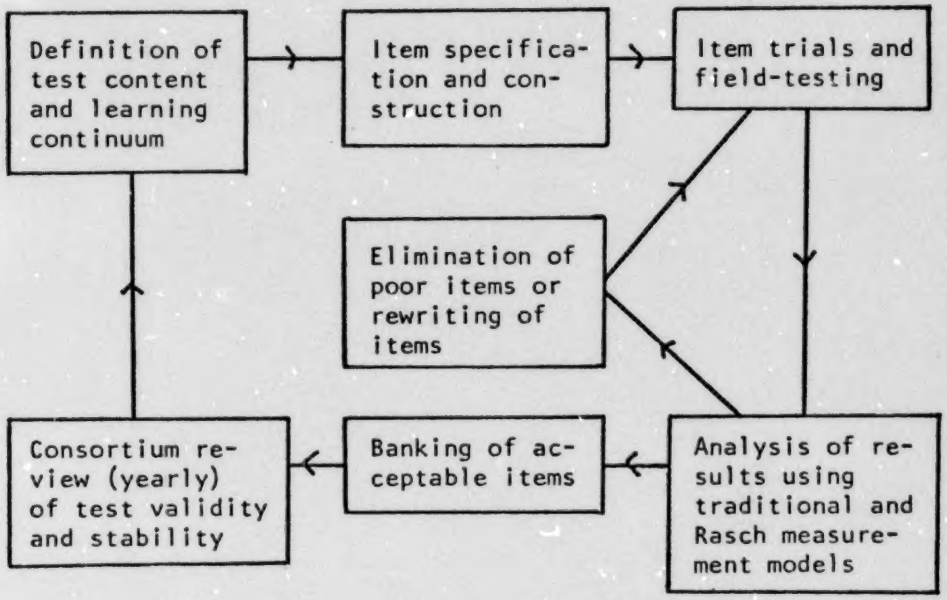
specialists who provided assurance that the test items would match life skills curricula and reflect the real-world criteria established for a competency-based life skills program. In terms of reliability, the internal consistency of the test items was determined through field testing that involved over 100,000 students; the ability, achievement and social characteristics of these students reflect those of a national norming sample. KR-20 reliability coefficient from these field tests ranged from .80 to .90, and were uniformly high across all groups tested. In addition to reporting on these data, the Manual provides instructions for the construction and administration of customized tests for various program uses.

Curriculum Index and Matrix. To link the assessment procedures with competency-based curricula, CASAS has created a curriculum index and matrix. The matrix identifies the location of each competency in a range of life skills instructional materials, and indicates the appropriate program level of these materials. The materials referenced in this matrix were those that met the curriculum standards established by the CASAS Consortium.

Implementation Workshops. To assist agencies in the implementation of CASAS, several workshops have been designed for agency staff. These workshops provide training over a three-day period, and address the areas of: a) test item writing; b) test construction using the CASAS Item Bank; c) designing and implementing an assessment system; d) classroom use and interpretation of test results; and e) Program Manager's use of data.

Development of an Item Bank

A design was established to provide for both internal and external validation of the CASAS item bank. As a formative or utilization-focused design, its path was cyclical leading to ultimate validation of items which were then banked into their present form or organization. The design is outlined below.


Definition of test content and learning continuum
Item specification and construction
Item trials and field-testing
Elimination of poor items or rewriting of items
Consortium review (yearly) of test validity and stability
Banking of acceptable items
Analysis of results using traditional and Rasch measurement models

Analysis results included computation of traditional statistics, including means, standard deviations, point-biserials, chi-square distributions, and P-values. Data were also screened according to Rasch Program parameters to include Rasch units (RITS) and statistics related to logistic curve analysis. The item-validation design was especially appropriate, as the normative or standardization model typically used for group tests does not fit the adult student population. Achievement of adult students along the dimension of life-skills did not prove to be normally distributed, and was better modeled by the logarhythmic curve of the Rasch model. The negatively skewed distribution of scores which are typically generated by criterion-referenced testing systems also held true for the adult population, providing additional evidence that the evaluation design was highly appropriate.

Validity

Content Validity is defined as the degree to which test items or instruments reflect the concrete or "real world" aspects of some criterion. Content and construct validity has been evaluated as the degree to which the CASAS items reflect the concrete aspects of criteria established for the demonstration of life-skills.

Initial work in the consortium began in 1980-81 with a group definition of a continuum of learning. This continuum is based upon a comprehensive and carefully defined taxonomy of adult life-skills compiled by adult educators representing approximately 30% of all federally funded adult education agencies in California. Within the general life-skills competency areas of Health, Occupational Knowledge, Government and Law, Consumer Economics and Community Resources, specific objectives were formulated and "mapped" or related to a continuum corresponding to beginning English-as-a-Second-Language instruction and continuing to intermediate-advanced levels through instruction for high school completion. The instructional outcomes or actual adult behaviors, expressed as competency statements (performance indicators), were generated and agreed upon as valid indicators for demonstration of student mastery or certification at each program level.

Adult educators with extensive experience and expertise agreed upon the content areas and objectives to be assessed at each program level. This process provided assurance that the design for measurement would match curriculum and reflect the real world criteria established for a competency-based life-skills program.

The code list includes well over 100 competency statements, which are keyed to the competencies and content areas. The competencies selected for assessment represent over 80% agreement among 40 consortium agencies as to relevance and instructional appropriateness. The statements have served as the basis for item construction.

A common concern shared by those who conduct research is the threat that faulty instrumentation presents to internal validity or integrity of experimental designs. Evidence presented throughout the remainder of this report, which includes a description of the process used to develop and validate items, confirms the appropriateness of using the item bank to measure adult life-skills achievement. The sensitivity, stability and validity of these items makes possible a variety of research projects on adult learning which are able to be carried out with unprecedented precision.

The compilation of content areas, competencies and statements provided a consistent format and starting point for item specifications. Tasks were defined to accommodate multiple-choice formats corresponding to competency statements. Potential item contributors received intensive training to include methods for writing acceptable stems and distractors, identifying displays and reviewing items. A skeleton of the item specifications follows.

All items measure a competency statement found in the CASAS Competency Code List as well as a specific Task (situation-specific-requirement).

A four-digit coding system is used to define each of the sample items. The first digit represents the CONTENT AREA. These areas are as follows:

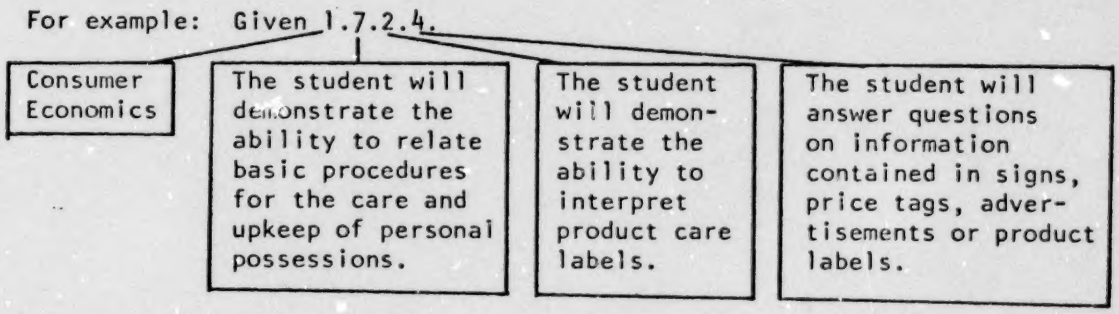
1 - Consumer Economics
2 - Community Resources
3 - Health
4 - Occupational Knowledge
5 - Government and Law

The second digit represents the COMPETENCY found within the designated CONTENT AREA.

The third digit represents a specific COMPETENCY STATEMENT.

The fourth digit represents the TASK (situation-specific-requirement) as follows:

1 - Completing fill-in-the-blank forms
2 - Answering questions based on information contained in consumer billings, maps, charts, matrices, graphs or tables
3 - Answering questions based on information contained in stories, articles, paragraphs, sentences, directions or pictures
4 - Answering questions on information contained in signs, price tags, advertisements or product labels
5 - Answering questions based on information contained in measurement scales or diagrams
6 - Selecting correct picture or aural response from oral cues


For example: Given 1.7.2.4.
Consumer Economics
The student will demonstrate the ability to relate basic procedures for the care and upkeep of personal possessions.
The student will demonstrate the ability to interpret product care labels.
The student will answer questions on information contained in signs, price tags, advertisements or product labels.

Item writers developed test items for each program level. Field trials were conducted to screen faulty items through in-depth teacher and student comment and criticism. Items were refined based on these critical findings. Once acceptable items had been written, formal field tests were formatted for extensive try-out and review beginning in the Fall of 1981. Due to the interim field-trial procedures soliciting teacher and student comments, very few items were rejected through the formal field-test results. Over 100,000 adult students have been involved in formal field testing that established the calibrated difficulties of the items.

Questions about validity ultimately translate into meaningfulness of results. In the preceding section, evidence has been presented to trace and describe a carefully defined process for assurance that the assessment items measure what they purport to measure.

Reliability

Evidence of reliability is presented in this section primarily as support for validity claims. The procedure used to validate and construct the CASAS item bank is described in relation to consistency of results obtained over time.

The general techniques applied for item validation included a combination of classical or traditional statistical methods and the more contemporary single parameter Rasch Measurement Model. Rasch analysis yields a difficulty index or calibration for each item analyzed. The most important feature of this type of analysis is the creation of an equal interval scale or continuum represented by items at low or beginning difficulty through advanced or high levels of difficulty. Each item is "fixed" or set on the continuum based on student performance during item trials.

The relative difficulty of test forms may also be established through computation of the average RITS (Rasch Units) -- where RIT = logit x 10 + 200. Since item and test difficulty are set through an analysis of the performance of students representing various program levels, it was important to establish a level of difficulty or typical performance in RITS at a specific program level. The performance of intermediate ESL students on field test Form 74 was used as a reference point or "base" from which to determine relative difficulty of all other test forms. A comparison of forms is made possible through an item-block design, which is to say that tests were formatted in accordance with the blueprint used to construct the base test. Item blocks made up of items that address common competency areas and statements across tests were used. While items vary from one form to another, some items were retained in parallel forms to preserve measurement characteristics and increase the odds that forms were indeed equivalent. In the Fall of 1981, a total of 21 tests were analyzed. Included in the 21 tests was the base or anchor test (Form 74), which includes 19,000 ABE and ESL students. Field trial performances on other forms were linked or related back to the base test. Linking values were assigned based on + or - the number of Rasch Units computed for the base test. The linking process has been outlined in this section to demonstrate how the curricular scale was created. Once test forms, and therefore items, were adjusted via linking to the base test, it became possible to describe groups of students (beginning through advanced) in terms of a mean and range of difficulty on a curricular scale from 160-250 RITS.

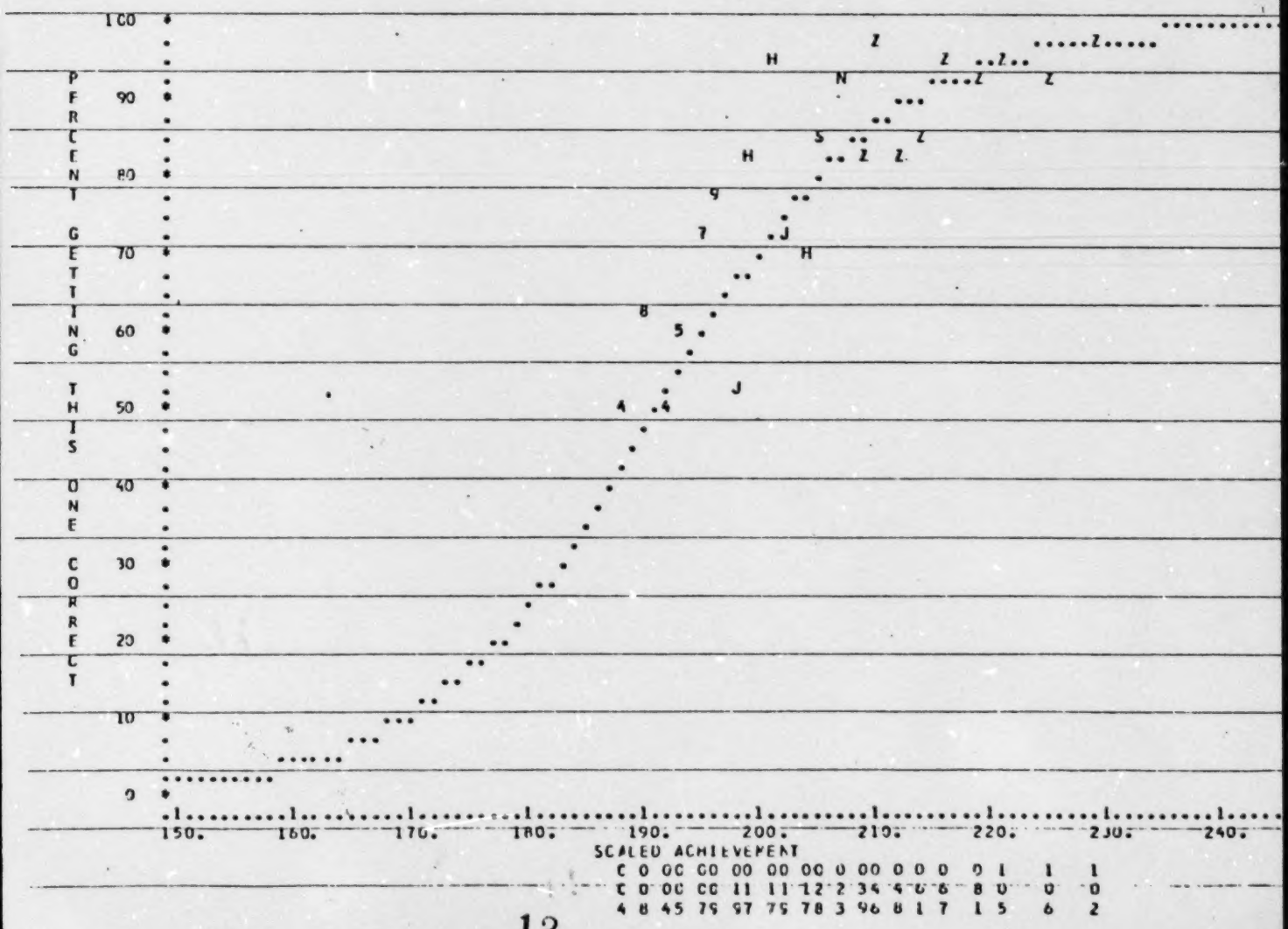
High-Low reports (top and bottom 27% of students) were used to analyze item-distractors; response patterns varying from those expected resulted in rewriting of items/distractors, elimination and/or additional field-testing and calibration. KR-20 correlation coefficients for the 21 tests ranged from .80 to .90. Rasch program parameters screen scores above 95% and below 20%. The logic of this procedure is sound when one considers that those who score very high or at chance levels contribute little information to the analysis and are not adequately measured. Other data provided by the Rasch program include the point and adjusted or non-spurious bi-serial correlation coefficients, item difficulty (RITS), standard error (RITS), mean square fits and refined fits. These statistics were used to interpret item behavior as a comparison of the item curve against the actual test curve. The following graph displays the characteristic curve for Item #1.

REPORT AM6R40 — SAN JOSE UNIFIED SCHOOL DISTRICT — PAGE 7
RASCH TESTING SYSTEM ITEM ANALYSIS — 5/ 5/82
C.A.S.A.S. ALL STUDENTS

TEST 4 FORM 74 GRADES = ALL

**** THE CHARACTERISTIC CURVE FOR ITEM 1 (241403) *****

RIT= 191.2 PCT= 88.8 PT.BIS=0.28 PR.SC.FIT= 1.06 REFINED FIT= 1.11 AVERAGE DEVIATION= 6.25
TOTAL NUMBER OF STUDENTS= 834



Numbers and letters represent plots of students achieving the same scaled achievement score on the test; plots far above the test curve display students scoring higher than expected, while plots far below the curve display those scoring below expectancy. Subjecting items to this type of graphic logistic curve analysis enables the evaluator to view the items' relationship to the total test curve. In the above case, the mean square fit is within acceptable limits. That is to say that performance on the item meets standards or ranges of expected performance. If mean square fits exceeded certain parameters (groups of students displayed extreme deviations from the curve), the item was questioned, revised, eliminated or rewritten and retested.

Items were eliminated, rewritten or recalibrated based on repeated trials with student populations at the same program level (beginning-advanced). The results have been consistent with the findings of Ingebo, Doherty and Davis (1982). Items included in the bank have consistently calibrated within a + or - 3 RITS difficulty over a three-year period.

Transportability, Generalizability

CASAS is transportable to and replicable in a wide variety of educational settings and student populations. This claim is made on the basis of educational settings and populations used during item validation. Initial item calibration was based upon analyses of responses from over 17,000 students for listening and reading life-skills field-tests in 1981-82. During 1982-83, new items and test forms were tested and stability of old items was checked with over 20,000 students enrolled in Adult Basic Education, English-as-a-Second-Language, and high school completion programs. The size and setting of programs varied from large, urban schools such as Los Angeles Unified (280,000 +) to small, rural schools such as Elsinore Valley (60 + students). Educational settings included correctional institutions, hospitals, unified school districts, community college districts and private, non-profit organizations. The following chart displays the results of Fall 1981 field-testing with B (intermediate) and C (advanced) students. Included is a breakdown of test means and related statistics by ethnicity.

C A S A S FALL'81 FIELD TESTING RASCH ITEM CALIBRATION AND TEST ANALYSIS

LEVEL/ FORM	RUN #	POP. TYP	N	36 REPORT RS	STD DEV	KR-20	LOW RANGE	HIGH RANGE	# OF ITEMS	N	48 REPORT RS	STD DEV	KR-20	LOW RANGE	HIGH RANGE	UNREL RIT
B/58	1	ALL STU	255	27.46	6.11	.862	4	37	37	245	27.35	5.68	.837	8	35	214.59
	2	SPANISH	74	27.97	5.5	.837	10	35	37	74	27.97	5.5	.837	—	—	———
	3	VIETNAM	63	27.35	5.69	.834	9	35	37	63	26.95	5.69	.834	—	—	———
	4	CHINESE	58	28.36	4.67	.788	13	37	37	57	28.21	4.57	.776	—	—	———
	5	INT ESL	238	27.13	6.11	.859	4	37	37	229	27.03	5.66	.832	8	35	213.86
C/62	58	ALL STU	220	31.42	8.39	.914	8	43	43	209	30.59	8.03	.9	9	48	212.7
	59	ENGLISH	38	32.26	7.07	.891	12	41	43	35	31.51	6.86	.879	—	—	———
	60	SPANISH	93	32.71	7.97	.912	8	42	43	82	31.95	7.41	.891	9	48	215.83
	61	VIETNAM	32	28.47	9.59	.932	9	43	43	30	27.57	9.22	.923	—	—	———
	62	INT ESL	25	28.4	7.53	.882	12	36	43	25	28.4	7.53	.882	—	—	———
	63	ADV ESL	99	33.08	7.85	.909	8	43	43	86	32.2	7.28	.885	9	48	215.29
	64	ABE	65	32.19	6.76	.869	12	42	43	60	31.42	6.47	.85	—	—	———

The KR-20 reliability coefficients are uniformly high across groups; the variance or dispersion of scores across groups is similar, and the raw score means are comparable. While the display might be used to demonstrate that the items do not offer any systematic bias against a specific group, it is included as an indication of generalizability to other, similar groups. Due to the magnitude of the sample, diversity in ethnicity, stability in performance and variety in educational setting, the items and tests are likely to perform in a similar manner with student populations in other states.

In summary, evidence of statistical reliability has been presented over time. The basis for claims of reliability has been established with groups of students qualifying as a large, national sample. Evidence has been presented that the circumstances under which item validation took place was sufficiently diverse (population and setting) to provide reasonable assurance of generalizability. CASAS has the potential for increasing the precision in placement, monitoring, certification and evaluation of students and programs at local, state and national levels. Items and tests are also educationally meaningful in relation to the instructional accountability of life-skills programs and internal validity of research on teaching methods and materials using life-skills achievement as the dependent variable.

REFERENCES

Bernardin, H. J., Behavioral expectation scales versus summated scales: A fairer comparison. Journal of Applied Psychology, 1977, 62 , 422-428.

Dunnette, M. D. and Borman, W. C., Personnel selection and classification systems. In L. W. Porter and M. R. Rosenzweig (Eds.), Annual Review of Psychology, 1979, 30, 477-525.

Haladyna, T. M. and Roid, G. H., "A Comparison of Two Approaches to Criterion-Referenced Test Construction," Journal of Educational Measurement Vol. 20 No. 3, Fall, 1983, 271 pages.

Hambleton, R. K. (Editor), Applications of Item Response Theory, Educational Research Institute of British Columbia, Vancouver, B. C., (1983), 233 pages.

Ingebo, G. S., Doherty, V. W. and Davis, J., "Stability of Item Calibrations in a Rasch-Based Testing Program," AERA Round Table Presentation, 1982, New York.

Jacobs, R., Kafry, D., and Zedeck, S., Expectations of behaviorally anchored rating scales. Personnel Psychology, 1980, 33, 595-640.

Kolen, M. S. and Whitney, D. R., "Comparison of Four Procedures for Equating the Tests of General Educational Development," Journal of Educational Measurement B-19 No. 4, Winter, 1982, pages 279-294.

Landy, F. and Farr, J. Performance rating. Psychological Bulletin, 1980, 87, 72-107.

Rasch, G., "Probablistic Models for Some Intelligence and Attainment Tests," Copenhagen, Denmark: Danmules Paedogogiske Institute, 1960.

Rentz, R. Robert, "Innovations in the Practice of Measurement," in New Directions for Testing and Measurement, Dale Carlson (editor). A Quarterly Sourcebook No. 10, Jossey-Bass, Inc., Publishers, San Francisco, (1981) pages 91-98.

Wright, B. D. and Stone, M. H., Best Test Design, Mesa Press Chicago, 1979.

Wright, B. D., "Sample-Free Test Calibration and Person Measurement," Proceedings of the 1967 Invitational Conference on testing problems. Princeton, N.S., Educational Testing Services, 1968.